The EMBO Journal vol.12 no.4 pp.1363–1374, 1993

Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells

Eric J.Bellefroid[1], Jean-Christophe Marine[1], Thomas Ried[2], Pierre J.Lecocq[1], Michele Rivière[3], Chris Amemiya[4], Dominique A.Poncelet[1], Pierre G.Coulie[5], Pieter de Jong[4], Claude Szpirer[3], David C.Ward[2] and Joseph A.Martial[1]

[1]Laboratoire de Biologie Moléculaire et de Génie Génétique, Université de Liège, Institut de Chimie B6, B-4000 Sart-Tilman, Belgium, [2]Department of Genetics, Yale University School of Medicine, New Haven, CT 06510, USA, [3]Département de Biologie Moléculaire, Université Libre de Bruxelles, B-1640 Rhode-St-Genèse, Belgium, [4]Lawrence Livermore National Laboratory, University of California, POB 5507, Livermore, CA 94551, USA and [5]Institut Ludwig, Université Catholique de Louvain, Avenue Hippocrate 7459, B-1200 Bruxelles, Belgium

Communicated by W.Fiers

KRAB zinc-finger proteins (KRAB-ZFPs) constitute a large subfamily of ZFPs of the *Krüppel* C_2H_2 type. KRAB (*Krüppel*-associated box) is an evolutionarily conserved protein domain found N-terminally with respect to the finger repeats. We report here the characterization of a particular subgroup of highly related human KRAB-ZFPs. ZNF91 is one representative of this subgroup and contains 35 contiguous finger repeats at its C-terminus. Three mRNA isoforms with sequence identity to ZNF91 were isolated by the polymerase chain reaction. These encode proteins with a KRAB domain present, partially deleted or absent. Five genomic fragments were characterized, each encoding part of a gene: the ZNF91 gene or one of four distinct, related KRAB-ZFP genes. All exhibit a common exon/intron organization with the variant zinc finger repeats organized in a single exon and the KRAB domain encoded by two separate exons. This positioning of introns supports the hypothesis that the mRNA isoforms encoding polypeptides with variability in the KRAB domain could arise by alternative splicing. By *in situ* chromosomal mapping studies and by analysis of fragments from a human genomic yeast artificial chromosome library containing KRAB-ZFP genes, we show that these genes occur in clusters; in particular, a gene complex containing over 40 genes has been identified in chromosomal region 19p12-p13.1. These ZNF91-related genes probably arose late during evolution since no homologous genes are detected in the mouse and rat genomes. Although the transcription of members of this KRAB-ZFP gene subgroup is detectable in all human tissues, their expression is significantly higher in human T lymphoid cells.

Key words: alternative splicing/gene cluster/human/T lymphocytes/zinc finger

Introduction

Eukaryotic nucleic acid binding proteins can be classified according to the conserved structural motifs they have in common. One such motif is the C_2H_2-type zinc-finger repeat discovered in *Xenopus* transcription factor IIIA (TFIIIA) and which can confer specific RNA-, DNA-, or RNA- and DNA-binding activity (Miller *et al.*, 1985, reviewed in El-Baradi and Pieler, 1991). This C_2H_2-type zinc-finger motif has since been found in a number of proteins with key developmental regulatory functions: in *Drosophila Krüppel* (Rosenberg *et al.*, 1986) and *hunchback* (Tautz *et al.*, 1987) ZFPs, involved in embryonic pattern formation; in the mouse Krox-20 zinc finger protein (ZFP), active in both patterning of the hindbrain and control of cell proliferation (Chavrier *et al.*, 1988; Wilkinson *et al.*, 1989); in ZFPs thought to be involved in tumorigenesis, including Gli-3, which is implicated in Greig syndrome (Vortkamp *et al.*, 1991) and WT1, which is altered in Wilms' tumors (Haber *et al.*, 1990).

Besides these well characterized ZFPs, vertebrate genomes encode a large number of ZFPs which have mostly been isolated on the basis of structural homology in the zinc finger region and whose biological function is unknown. In humans, it has been estimated that there are 300–700 such different ZFP genes (Bellefroid *et al.*, 1989; Crossley and Little, 1991) and similar high numbers have been found in *Xenopus* (Köster *et al.*, 1988) and mouse (Chowdhury *et al.*, 1987). The vast majority of ZFPs are classified as *Krüppel*-like on the basis of the fact that they share a typical *Krüppel* feature, namely a conserved stretch of six amino acids (the H/C link) connecting successive finger repeats (Schuh *et al.*, 1986). Most of these ZFP genes encode multiple (up to 37 in the *Xenopus* Xfin ZFP) tandem repeats of the zinc finger unit, often organized in a single exon. Some of these ZFP genes, such as the mouse *Evi-1* (Bordereaux *et al.*, 1990) or Zfp-35 (Cunliffe *et al.*, 1990) genes, have been shown to produce primary transcripts subject to alternative splicing events giving rise to different proteins, thus further increasing the number of distinct ZFPs expressed in vertebrates. Several of these ZFP genes have been mapped to the same region, suggesting a clustered organization; a number of ZFP genes, deleted in the lethal mouse embryonic mutation t^{w18}, appear to be clustered on mouse chromosome 17 (Crossley *et al.*, 1991) and assignment of a large number of ZFPs to their chromosomal locations provides further evidence of clustering (Huebner *et al.*, 1991; Hoovers *et al.*, 1992; Lichter *et al.*, 1992; Rousseau-Merck *et al.*, 1992). Furthermore, two closely related *Drosophila* ZFP genes, termed *sry* α and β, are located within the same 6 kb genomic DNA fragment (Vincent *et al.*, 1985). Likewise, two closely related *Xenopus* ZFPs, XFG5–1 and XFG5–2, are encoded within a region only 8 kb long (Köster *et al.*, 1991).

Subfamilies of related ZFPs have been defined on the basis of conserved structural features found outside the finger domain. Such elements include the FAX domain found in a large number of *Xenopus* ZFPs (Knöchel *et al.*, 1989) and the evolutionarily conserved KRAB (*Krüppel*-associated box) domain identified in human, mouse and *Xenopus* ZFPs (Bellefroid *et al.*, 1991; Rosati *et al.*, 1991). The function of these conserved modules remains to be elucidated.

In this report, we define and characterize a particular subgroup of highly related human KRAB-ZFP genes. Using the ZNF91 gene as a model for this KRAB-ZFP subgroup, we have: (i) determined the sequence of the corresponding cDNA and identified by PCR three distinct mRNA isoforms differing in their coding capacity with respect to the N-terminal KRAB domain; (ii) characterized five genomic fragments, each encoding part of a ZNF91-like gene, either ZNF91 itself or one of four different related KRAB-ZFP genes exhibiting a common exon/intron organization; (iii) demonstrated, by isolation and analysis of yeast artificial chromosome (YAC) genomic fragments, the existence of multiple ZNF91-related genes clustered on chromosome 19p12-p13.1, (iv) showed that these ZNF91-related genes have a broad expression profile, the highest expression level being observed in T lymphoid cells.

Results

cDNA and predicted amino acid sequence of a human KRAB-ZFP, ZNF91

A human placental cDNA library and a human genomic library were screened with an oligonucleotide probe corresponding to the consensus H/C link sequence of the *Drosophila Krüppel* ZFP. This led to our identifying several hundred distinct ZFP genes (Bellefroid *et al.*, 1989). About one-third of these human ZFPs share, in their most N-terminal part, a very conserved region about 75 amino acids long, the KRAB domain (Bellefroid *et al.*, 1991).

To study these KRAB-ZFPs further, we characterized one cDNA encoding a representative of this family: ZNF91. This cDNA, isolated in a second round of screening of a human teratocarcinoma cell (NT2D1) cDNA library, was a longer copy of previously isolated cDNAs called HPF7 and HTF10 (Bellefroid *et al.*, 1989, 1991). From its sequence we predicted an open reading frame (ORF) coding for a 1182 amino acid protein if the second methionine (Figure 1, position 10), highly conserved in all KRAB-ZFP sequences (Bellefroid *et al.*, 1991) and best fitting the consensus sequence A/GCCATGG (Kozak, 1987), is taken to be the translation start site. The 171 nucleotide (nt) 3′ untranslated region contains a repetitive Alu element and is terminated by a poly(A) tail. No canonical AATAAA polyadenylation signal was found.

The predicted protein shows all the structural features which define the KRAB-ZFP subfamily. It contains a C-terminal region with many (35) contiguous finger repeats. These fingers, connected by H/C-link-type sequences, exhibit no unusual features as compared with the other human ZFPs (Bellefroid *et al.*, 1989), except that two groups of seven fingers (fingers 6–12 and fingers 17–23) show >97% amino acid sequence homology. In these fingers, all putative DNA-binding residues (Pavletich and Pabo, 1991) are conserved. This suggests that, if ZNF91 ZFP does indeed bind nucleic acids, its target may be a tandemly repeated sequence.

The N-terminal KRAB domain includes two modules, residues 11–52 (KRAB A) and residues 53–84 (KRAB B). These modules were first defined arbitrarily due to their absence or presence in individual KRAB-ZFP cDNA sequences (ZNF85 and ZNF92 cDNAs lack element B, Bellefroid *et al.*, 1991). Subsequently, the limits of ZNF2 (Rosati *et al.*, 1991) and ZNF43 KRAB-ZFPs (Lovering and Trowsdale, 1991) were found to correspond to points of divergence in alternative cDNA forms. Seeking such alternative transcripts of the ZNF91 gene, we synthesized two primers from the ZNF91 cDNA sequence encompassing the KRAB region and used them in PCR experiments to amplify mRNA fragments from various cell lines (Figure 2A). The 5′ primer was chosen in the variable 5′ non-coding region. The 3′ primer corresponds to the non-canonical first finger unit, while the specificity of these primers was first tested by using them to screen, by PCR, a total human genomic YAC library; from the total library, only five positive YACs were isolated; these were subsequently found to be derived from the same locus and to contain the ZNF91 gene (Table I). With total RNA from all cell lines examined (including MOLT-4, NT2D1 and THP-1), three distinct bands were identified. The DNA fragments were purified and sequenced. The major fragment is of the expected size (661 bp) and corresponds exactly to the sequence of ZNF91 cDNA. Two other fragments, 565 and 438 bp long, were also found to match the ZNF91 cDNA exactly, except for deletions. They most probably constitute two less abundant mRNA isoforms. Isoform α contains a 96 nt/32 aa deletion corresponding exactly to the B box of the KRAB domain and maintains the predicted ORF. Isoform β has an internal deletion of 223 nt and lacks the entire A and B boxes of the KRAB domain. This results in an ORF with the first initiation codon located at position 392 (Figure 2b).

The N-terminal KRAB domain of ZNF91 ZFP is connected to the finger cluster through a 92 aa spacer element which is 76–97% similar to that found in ZNF43, ZNF85, ZNF90, ZNF92 and ZNF93 (Bellefroid *et al.*, 1991). However, this spacer region, which is rich in cysteine and histidine residues not arranged in typical C_2H_2 fashion, differs totally from the spacer regions of the other known KRAB-ZFPs: ZNF7, ZNF8 (Lania *et al.*, 1990), ZNF10, ZNF15 (Thiesen *et al.*, 1991), Zfpl, Zfp2 (Chowdhury *et al.*, 1988, 1989), Xfin (Ruiz i Altaba *et al.*, 1987), ZNF2.2, ZNF41.1, HZF3.1, HZF7.1 (Rosati *et al.*, 1991) and ZNF45 (Constantinou-Deltas *et al.*, 1992). This spacer region, specific to our collection of KRAB-ZFPs, thus defines a special subset of KRAB-ZFPs.

Chromosomal location and structural characterization of ZNF91 and related KRAB-ZFP genes

Due to the exceptionally high degree of homology between ZNF91 and the other previously isolated KRAB-ZFP-encoding cDNAs (ZNF85, ZNF43, ZNF90, ZNF92 and ZNF93) (Bellefroid *et al.*, 1991), we speculated that the corresponding genes, probably generated by duplication, might be organized in a cluster. Therefore, we first determined their chromosomal assignment using the corresponding cDNA probes in Southern blot analysis of DNA from a human–rodent hybrid panel (Wathelet *et al.*, 1988). In contrast to another related KRAB-ZFP gene (ZNF117) previously mapped to chromosome 7q11.2 (Bellefroid *et al.*, 1992), all six loci were found on

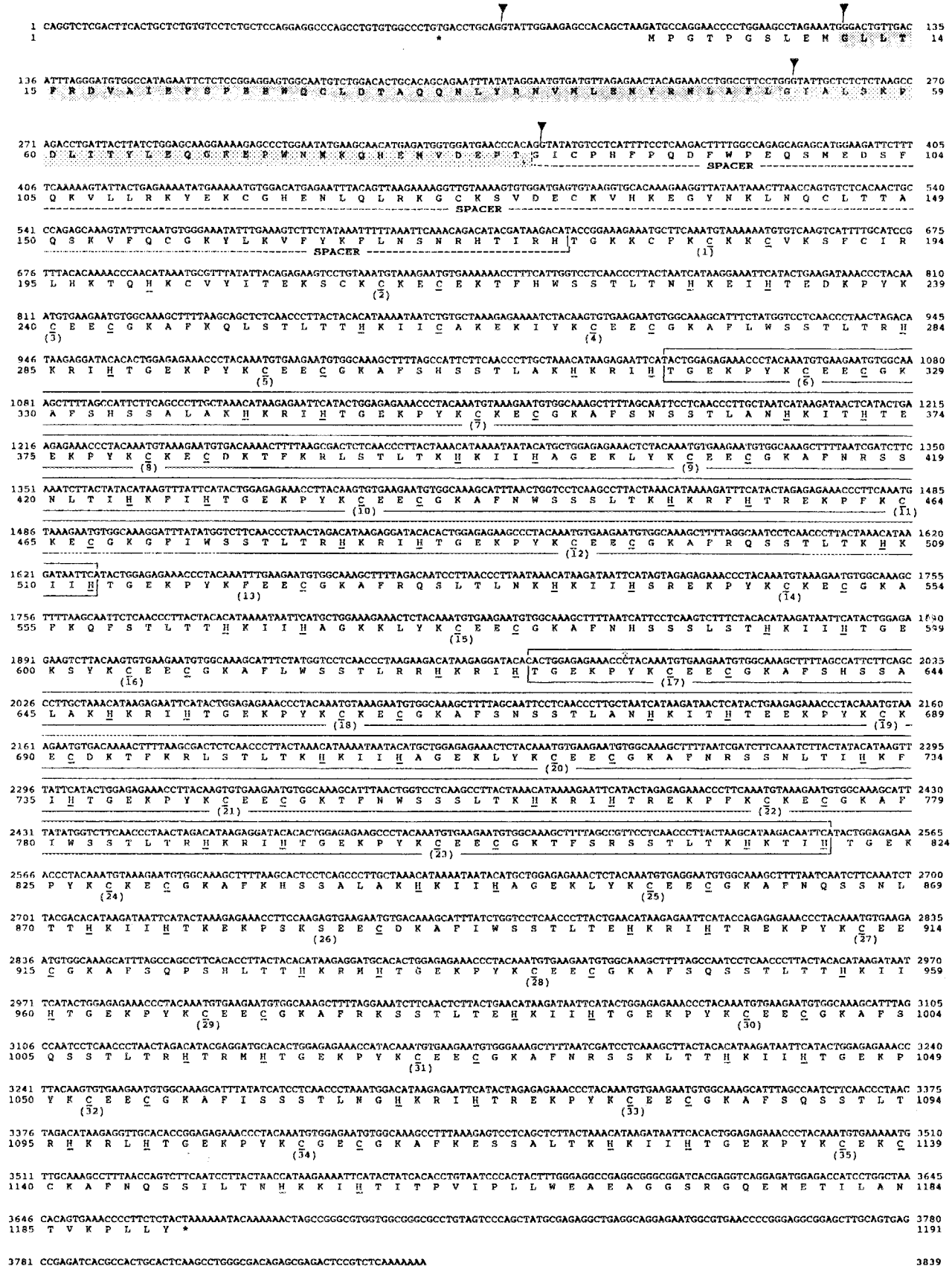

Fig. 1. ZNF91 cDNA and predicted protein sequences. (EMBL accession no. L11672). Arrowheads denote the positions of introns. The position of the first intron was predicted by sequence alignment with the corresponding sequences in the F31854 and F33152 genomic fragments. The 35 finger motifs are numbered. Within each finger motif, the conserved cysteine and histidine pairs are underlined. The seven reiterated fingers are boxed. The shaded areas correspond to the KRAB A and B elements. The spacer region connecting the KRAB and finger domains is indicated. In-frame stop codons are marked by (*).

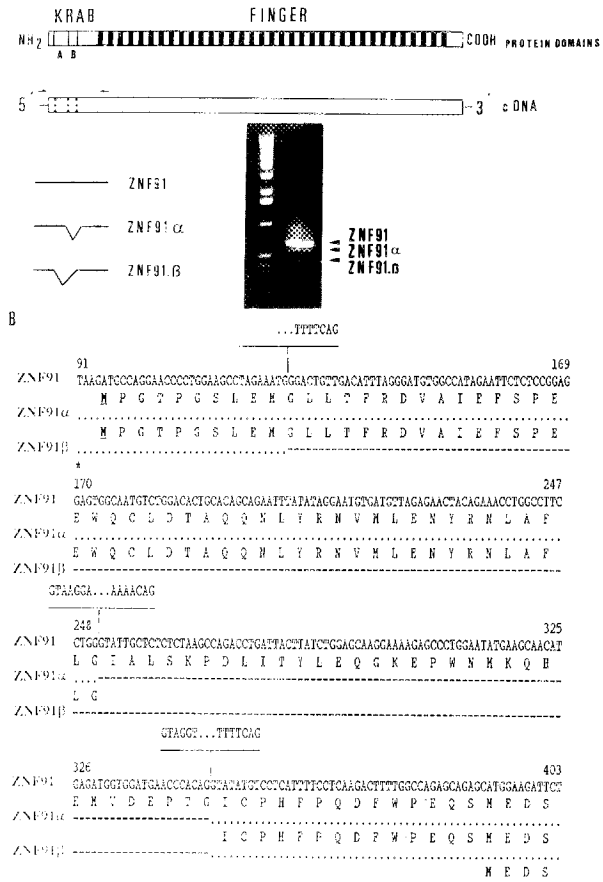

Fig. 2. Schematic representation of the structural domains of ZNF91 KRAB-ZFP and nucleotide and deduced amino acid sequences of alternative cDNA forms. (**A**) The ZNF91 protein structure is correlated with the exon/intron organization of the corresponding transcription unit. The pair of primers used for PCR analysis is indicated. The ethidium-bromide-stained agarose gel with the PCR products and the alternative RNA structures deduced from the nucleotide sequence analysis are shown [left lane, 1 kb ladder (BRL); right lane, PCR products; arrowheads indicate the positions of faintly stained fragments]. (**B**) Nucleotide sequences of alternative ZNF91 cDNA isoforms. Dots and dashes respectively denote identities and deletions. Conceptual translations are reported below each sequence. The first in-frame ATG found in each sequence is underlined. Numbers on the left refer to the ZNF91 nucleotide sequence as in Figure 1. The positions and sequence termini of introns as determined from the F21946 genomic fragment are reported. Asterisks denote in-frame translation termination codons.

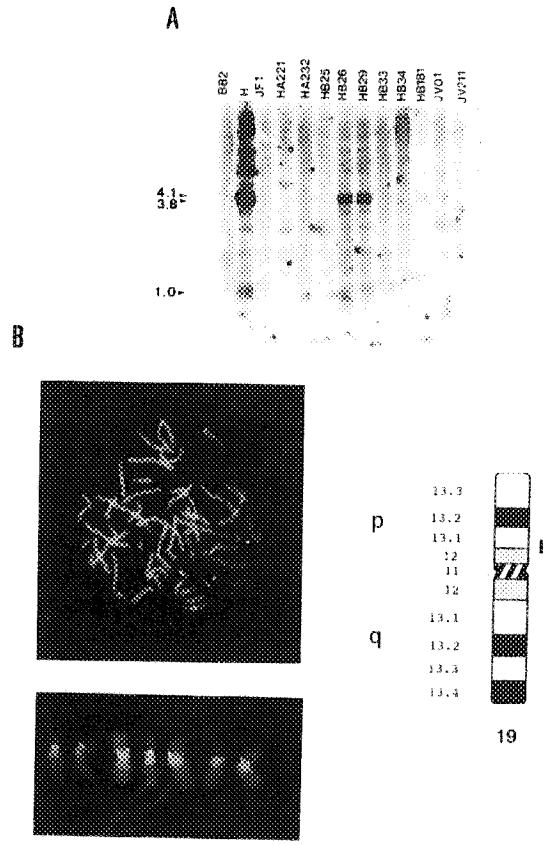

Fig. 3. Chromosomal localization of ZNF91 KRAB-ZFP gene. (**A**) Chromosomal assignment by Southern analysis of human–rodent hybrid cell DNAs. A Southern blot of DNA from mouse (B82), human (H), rat (JF1) and eight hybrid cell clones cleaved with *Eco*RI and hybridized with the ZNF91 cDNA probe. Three restriction fragments, detected at 1, 3.8 and 4.1 kb, co-segregate with human chromosome 19. For a description of the cell lines used, see Wathelet *et al.* (1988). (**B**) Regional mapping by fluorescence *in situ* hybridization. (Top) A human metaphase spread with the hybridization signals observed close to the centromere on the short arm of chromosome 19 (arrowheads). (Bottom) Composite of seven chromosome 19 homologs. On the ideogram of chromosome 19, the bar indicates the location of the ZNF91 gene according to the fractional length measurements.

chromosome 19. Fluorescence *in situ* hybridization with human metaphase chromosomes revealed that all these loci correspond to the short arm of chromosome 19, either the proximal part of band 19p13 or the distal portion of band 19p12. A representative blot and selected *in situ* hybridizations obtained with the ZNF91 cDNA probe are shown in Figure 3. Within the limits of spatial resolution afforded by the fluorescence *in situ* hybridization technique, it would appear that all six KRAB-ZFP genes map within a few mega-basepairs. These data were further confirmed by the gene mapping approach described by Lichter *et al.* (1990). The fractional length (FLpter) distance of the signal to the p arm terminus relative to the total chromosome length was found to be between 0.32 and 0.35; these values are consistent with an assignment of these genes to the interface of bands 19p12 and 19p13.

To isolate genomic fragments containing adjacent KRAB-ZFP transcription units, we screened a cosmid library specific for human chromosome 19 (de Jong *et al.*, 1989). A total of 10^4 cosmids were screened and 384 H/C-link-positive clones were pooled. This pool was first screened with a fragment of ZNF91 cDNA encompassing the conserved KRAB domain. When this probe was used under reduced stringency conditions, 24% of the ZFP-encoding cosmids gave a positive signal, indicating that about one-quarter of the ZFP genes located on chromosome 19 are members of the KRAB family (Figure 4). Half of them are closely related to ZNF91 since under the same reduced stringency conditions these clones also gave a signal with a ZNF91 cDNA probe corresponding to the spacer region (Figure 4). The cosmid array was also rehybridized under high stringency conditions with probes derived from the finger region, which constitutes the most variable segment, from the different ZNF91 family members (i.e. ZNF91, Figure 4). The results indicated that, on average, each distinct ZNF91-related KRAB-ZFP gene is represented by only 2.5% of the spacer-positive cosmids and 1.2% of the KRAB-positive cosmids. This suggests that chromosome 19 is particularly rich in members of the KRAB subfamily. It probably contains ~80 members, some 40 of which are

Table I. Characteristics of the YACs containing ZNF91-related genes

YAC	Size (kb)	Gene Content	Chrom. Loc.	ZNF genes																									
				95.	43.	91.	94.	106.	122.	119.	90.	93.	110.	103.	104.	102.	109.	112.	105.	111.	85.	120.	100.	101.	118.	113.	117.	107.	108.
71D9	630	5	19p12-p13.1/8q23	+	+																								
179B7	240	2	19p12-p13.1			+																							
349F6	290	4	19p12-p13.1/13q21			+																							
245F12	500	7	19p12-p13.1		+	+																							
143B11	400	6	19p12-p13.1			+	+																						
140D11 *	370	nd	19p12-p13.1			+	+	+																					
414H7	260	2	19p12-p13.1				+	+																					
60D10	nd	7	19p12-p13.1				+		+																				
138D1	660	6	19p12-p13.1						+	+	+																		
332C9	620	6	19p12-p13.1									+	+																
158C9	490	6	19p12-p13.1											+	+														
335H4	380	4	19p12-p13.1													+	+												
84A10	490	3	19p12-p13.1													+													
131F11	220	4	19p12-p13.1													+	+	+											
88C4	370	8	19p12-p13.1															+	+	+	+								
11B6 *	nd	nd	19p12-p13.1																	+		+							
96A1	430	7	19p12-p13.1																	+	+								
140G4	380	4	19p12-p13.1																		+		+						
143B6	430	2	19p12-p13.1																					+					
143D6	500	2	19p12-p13.1																					+					
16B10 *	500	nd	19p12-p13.1																					+					
411H1	490	4	19p12-p13.1/13q21																						+	+			
63G11	450	2	7q11.2																								+		
172H5	480	3	7q11.2																								+	+	
176H5	490	1	3q13																										+
176G8	490	1	3q13																										+

YACs are ordered as deduced from the data in Figures 6–8 assuming they are not chimeric. YAC sizes were determined by PFGE of undigested YAC DNAs. The estimated number of ZNF91-related genes present in each YAC is deduced from the number of signals observed in Figure 6 and other autoradiograms, assuming the absence of restriction sites corresponding to the enzyme used to digest the YAC DNAs within the spacer region. The distinct genes identified in each YAC by PCR and sequencing are indicated by (+). Chromosomal localization of the YACs was done by *in situ* hybridization with Alu-PCR products. (* : YACs showing some degree of instability; nd: not determined)

closely related to ZNF91. This is in good agreement with our sequence analysis of the spacer region of 12 positive cosmids, among which eight distinct cosmids were found (F11465/ZNF94, F13463/ZNF95, F15034/ZNF96, F6644/ZNF97, F7175/ZNF98 and F8281/ZNF99), with F12460 and F21946 corresponding respectively to the previously isolated ZNF93 and ZNF91 cDNAs (Figure 7).

Figure 5 shows the structure of five genomic clones. One contains part of the ZNF91 gene and the others parts of four distinct related KRAB-ZFP genes. These cosmids, termed F11465, F21946, F28423, F31854 and F33152, were isolated by means of cDNA probes representing the entire ZNF91 mRNA (5′ untranslated region, KRAB, spacer and finger region). Exons containing fragments of individual cosmids were identified by hybridization of restriction enzyme digests with parts of ZNF91 cDNA. The sequences along the exons and at exon/intron junctions were determined on appropriate genomic subclones with the help of exon-specific oligonucleotide primers.

Cosmid F21946 contains regions whose sequences perfectly match portions of ZNF91 cDNA. Three regions correspond with putative exons. The A and B boxes of the KRAB domain are organized in separate exons (exons 3 and 4); in contrast, the entire finger domain, the preceding spacer region and the 3′ untranslated region are encoded by a single exon (exon 5). There is no sequence homology with the region upstream from the KRAB domain in ZNF91 cDNA. All homologous cDNA segments are flanked by the AG and GT splice junction sequences and, in most cases, these regions are in good agreement with the complete consensus splice junction sequences (Shapiro and Senapathy, 1987). The positions of these exon/intron junctions exactly correspond to the points of divergence of the ZNF91, ZNF91α and ZNF91β cDNA isoforms (Figure 2). The differences between these isoforms are thus compatible with the occurrence of alternative splicing with the primary transcript of the ZNF91 gene.

The four other genomic fragments used carry parts of distinct ZNF91-related KRAB-ZFP genes. As no matching cDNAs were isolated, the predicted exon/intron organization was deduced by aligning the genomic sequences with the ZNF91 cDNA sequence and identifying consensus splice donor and acceptor sites.

F11465 carries regions homologous to the ZNF91 cDNA which correspond to exons 4 and 5 of a KRAB-ZFP gene, called ZNF94, located near the ZNF91 gene (Figure 7 and Table I). Cosmid F31854, like F21946, includes regions corresponding to exons 3, 4 and 5 of another ZNF91-related KRAB-ZFP gene named ZNF114 (Figure 7). It contains another region homologous to the ZNF91 cDNA (83% homology) corresponding to a putative 53 bp exon (exon 2) encoding the 10 amino acids immediately preceding the

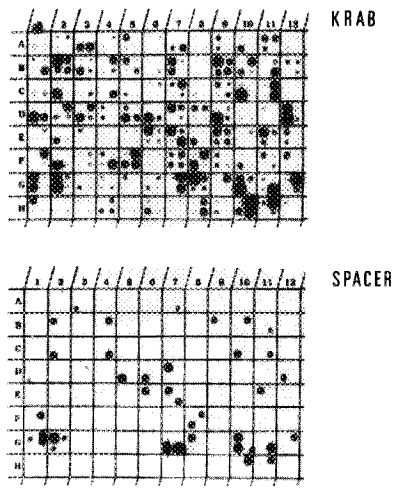

Fig. 4. Hybridization of ^{32}P-labeled KRAB or spacer probes from ZNF91 cDNA to dot-blot grids of 384 chromosome-19-specific cosmid clones hybridizing with an H/C-link degenerate oligonucleotide probe. Arrows point to clones corresponding to the ZNF91 gene.

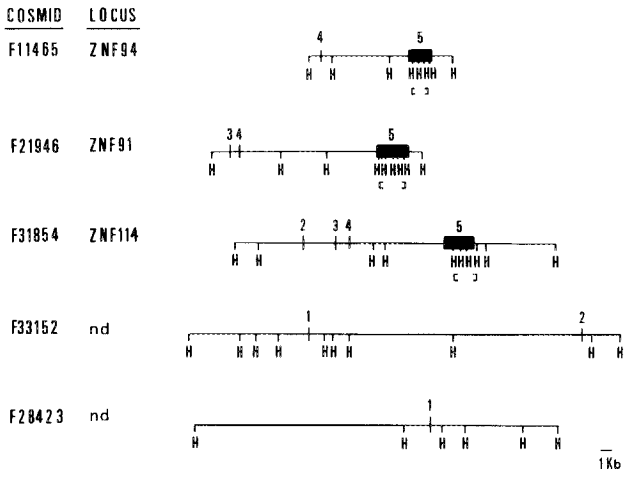

Fig. 5. Structural organization of five ZNF91-related, KRAB-ZFP-encoding genomic fragments. Exons are boxed and numbered. Brackets below the map indicate restriction fragments not ordered with respect to one another. *Hin*dIII restriction sites (H) are indicated. (nd: not determined).

KRAB domain and includes the 23 nt furthest 3′ in the upstream non-coding region. Cosmid F33152 also contains a putative exon 2 from another undetermined ZNF91-related gene. Twenty-two kilobases upstream of this exon, a second region of homology (85%) was identified. This aligns with the furthest 5′ part of the upstream non-coding region of ZNF91 cDNA. We assume that this upstream region corresponds to the first exon. Cosmid F28423 contains part (exon 1) of an additional unidentified KRAB-ZFP gene located in the 88C4 and 96A1 YAC clones by hybridization to YAC DNAs separated by PFGE (pulsed field gel electrophoresis). No other regions of homology with the ZNF91 cDNA probe were identified in the 30 kb genomic fragment.

Clustered organization of ZNF91-related KRAB-ZFP genes

None of the 58 isolated cosmids containing ZNF91-related genes characterized by Southern analysis seems to contain

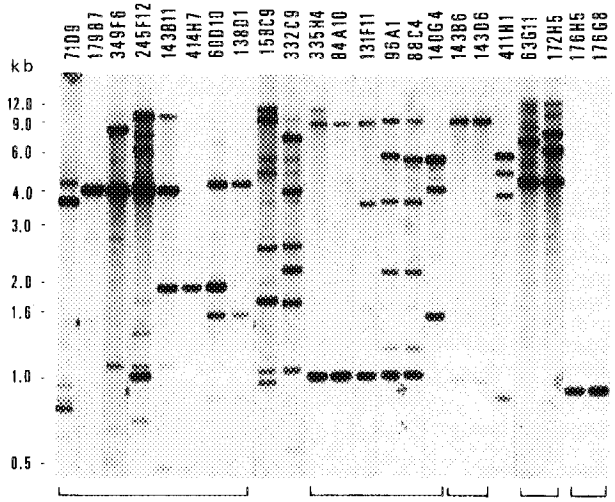

Fig. 6. Southern analysis of the YACs containing ZNF91-related genes. *Hin*dIII digests of the YAC DNAs were hybridized with the spacer probe derived from ZNF91 cDNA. Brackets indicate YACs with common hybridizing fragments. YACs are named according to the standard nomenclature of the Centre d'Etude du Polymorphisme Humain (CEPH). The 2.8 kb hybridizing band detected in each YAC is due to cross-hybridization with yeast DNA fragments corresponding to a repetitive sequence.

two adjacent KRAB-ZFP transcription units (Figure 5 and data not shown). To obtain genomic fragments with linked ZNF91-related genes, we therefore screened a human genomic YAC library (Albertsen *et al.*, 1990) by PCR. The degenerate primers used (see Materials and methods) amplify a 249 bp fragment of the spacer regions of the ZNF91 and related KRAB-ZFP genes. Twenty-six positive clones were isolated out of a total 20 000 screened clones. The positive clones contain YACs ranging in size from 240 to 660 kb, as determined by PFGE and Southern analysis with the spacer probe derived from ZNF91 cDNA (Table I).

Chromosomal localization of these YACs was done by fluorescence *in situ* hybridization to human metaphase chromosomes (Lengauer *et al.*, 1992) (Table I) The probes used were Alu-PCR products derived from the YAC DNAs. Except for YACs 71D9, 349F6 and 411H1, all other clones map to a single locus and thus appear to be non-chimeric. The two YACs bearing ZNF117 (see below, Table I) mapped to chromosome band 7q11.2, confirming our previous localization of this gene (Bellefroid *et al.*, 1992). YACs 176H5 and 176G8, which contain the same single ZNF91-related gene (ZNF108), were assigned to chromosome band 3q13. All the other YACs, as expected, mapped to chromosome band 19p12-p13.1.

To determine whether these YACs contain a single or multiple ZNF91-related genes, YAC DNAs were digested with *Hin*dIII or *Eco*RI, separated by conventional electrophoresis, and probed under low stringency conditions with the spacer fragment of ZNF91 cDNA (N372–N693). As shown in Figure 6, most YAC DNAs digested with *Hin*dIII gave multiple signals, indicating that they contain several (up to eight in 88C4) ZNF91-related transcription units. Bands of identical size were visible in several YACs, indicating that they may overlap (i.e, 179B7, 349F6, 2145F12, 143B11, 414H7, 60D10 and 138D1) and contain common ZNF91-related genes.

To establish further the relationships between the 26 YACs and to identify the KRAB-ZFP genes, we subcloned the

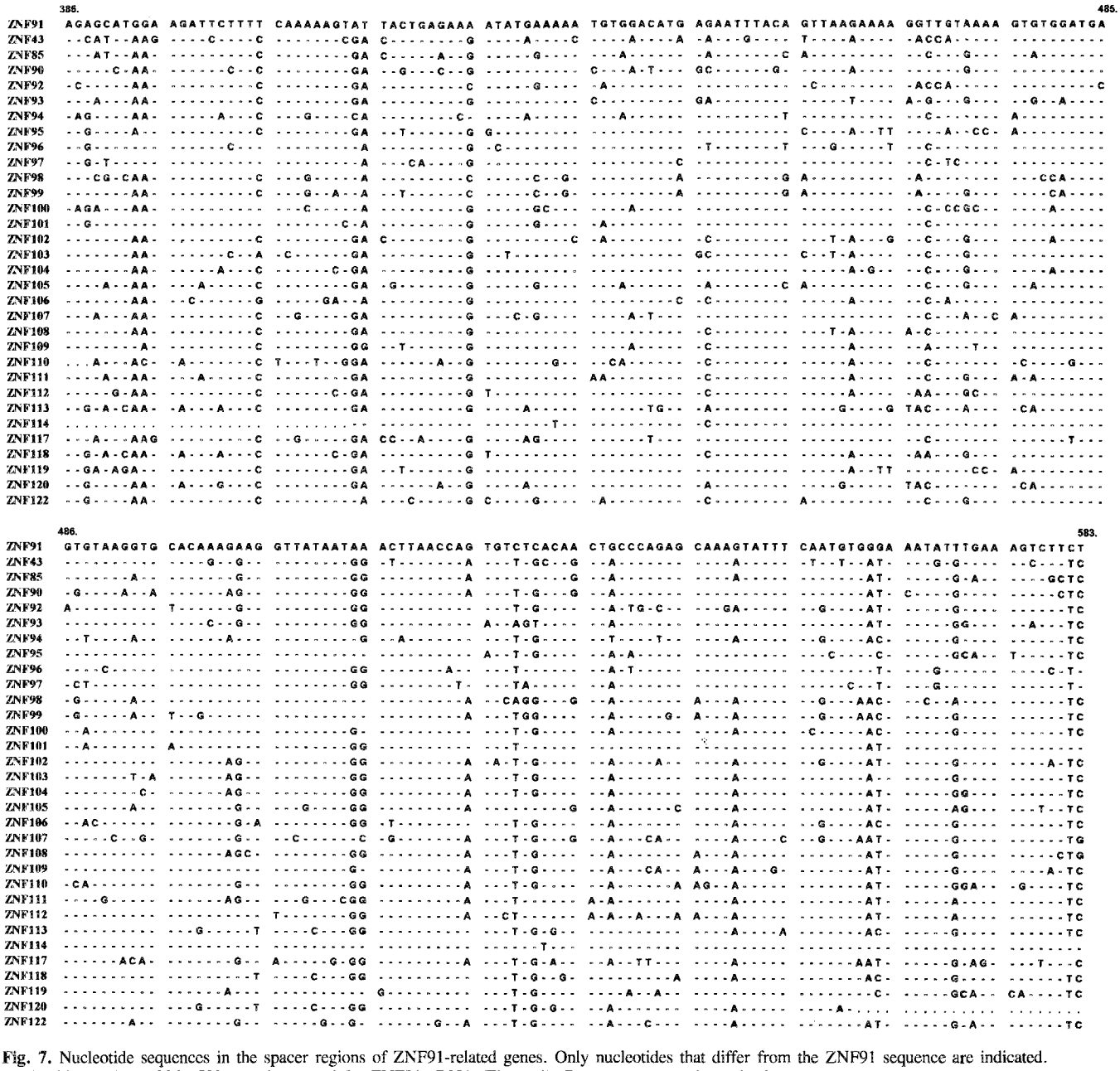

Fig. 7. Nucleotide sequences in the spacer regions of ZNF91-related genes. Only nucleotides that differ from the ZNF91 sequence are indicated. Nucleotide numbers (386–583) are those used for ZNF91 cDNA (Figure 1). Dots represent undetermined sequence.

249 bp PCR products for each YAC and sequenced 12 subcloned DNA fragments per YAC. Figure 7 shows the sequences derived from the distinct genes identified by this approach. When we compare these results with the estimated KRAB-ZFP content of these YACs (Table I), it is clear that we have not identified all the KRAB genes present in each YAC. However, on the basis of those genes which were detected, it is possible to identify YACs with overlapping inserts and to build contigs (Table I). The degree of amino acid similarity to the spacer region of ZNF91 ranges between 74 and 97%. These sequences represent 18 novel ZNF91-related KRAB-ZFP genes. Six out of eight previously described ZNF91 and related KRAB-ZFP cDNAs (Bellefroid *et al.*, 1991) were also represented in this set of YACs. Of the eight ZNF91 and related KRAB-ZFP genes isolated from the chromosome-19-specific cosmid library and for which sequence data are available in the spacer region (as outlined above), only three (ZNF91, ZNF94 and ZNF95) were re-

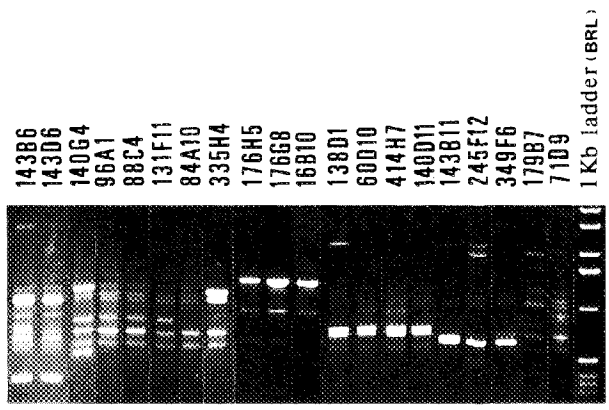

Fig. 8. Inter-Alu PCR products from the YAC DNAs resolved on a 1% ethidium-bromide-stained agarose gel. The Alu-repeat primer was TC-65 (Nelson *et al.*, 1989).

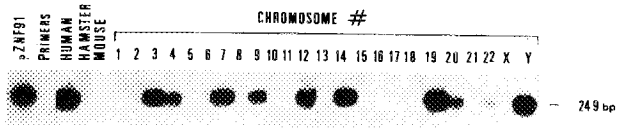

Fig. 9. Chromosomal mapping by PCR of the ZNF91-related genes. DNA samples were from the monochromosomal hybrid panel 2 (Coriell Cell Repositories). Degenerate primers were as described under Materials and methods. PCR products were analyzed by Southern blotting with the spacer probe from the ZNF91 cDNA. Control templates include ZNF91 cDNA and human, mouse and hamster genomic DNAs.

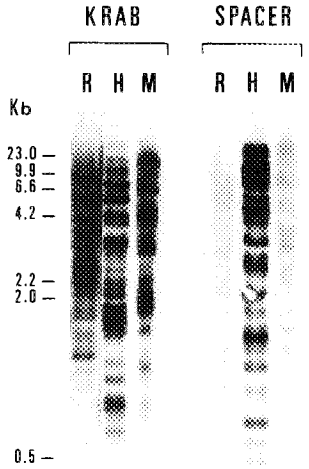

Fig. 10. Evolutionary conservation of KRAB- and spacer-encoding sequences. Southern blots containing *Eco*RI digests of human, mouse and rat DNA were hybridized under reduced stringency conditions with a KRAB (left) or spacer (right) probe derived from ZNF91 cDNA. kb, kilobase(s).

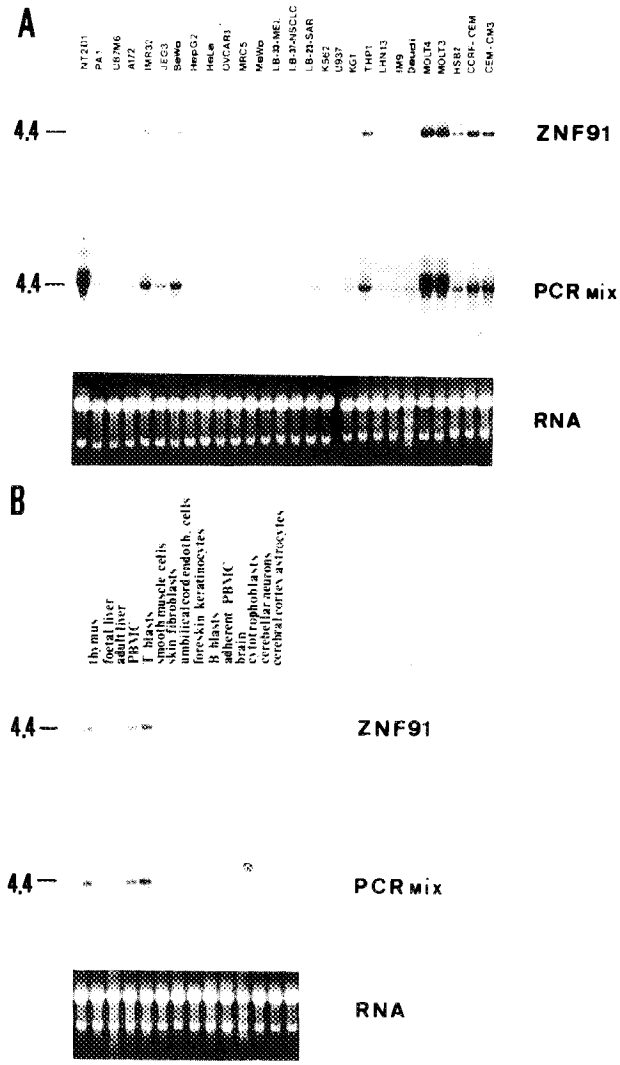

Fig. 11. Expression of ZNF91 and related genes in cell lines (**A**) and in tissues or normal cell populations (**B**). The probes were a 321 bp fragment of ZNF91 cDNA (spacer region) (top panels) or a 249 bp PCR random probe designed to detect all ZNF91 family members (lower panels). Equivalent amounts of RNA were loaded as determined by both spectrophotometry and ethidium bromide staining of the blotted RNA (Kroczek, 1989).

isolated in the YACs. Relationships between these cosmids and the YACs were confirmed by rehybridization of the filters containing *Hin*dIII digests of the YAC DNAs with *Hin*dIII fragments of the cosmids containing the ZNF91-homologous spacer region. These hybridizations also confirmed that the restriction fragments detected in the YAC DNAs are the same size as those detected in the cosmid DNAs (data not shown).

To confirm the YAC overlaps identified by their ZNF91-related gene content, we also used the Alu-PCR fingerprinting method (Nelson *et al.*, 1989). As shown in Figure 8, several YAC clones predicted to overlap display common Alu-PCR fragments.

As our results indicate the presence of ZNF91-related genes on chromosomes other than 7 and 19, we determined the chromosomal distribution of such genes. To this end, we performed a PCR assay, using the same degenerate primers used previously for YAC screening (see Materials and methods). The assay was done on DNA from a panel of monochromosomal human–rodent somatic cell hybrids representing each of the 24 different human chromosomes. Products of the PCR reactions were separated by electrophoresis, blotted and hybridized with a ^{32}P-labeled ZNF91 cDNA spacer probe to confirm the identity of the amplified DNA. Figure 9 shows that there was no amplification of the 249 bp fragment of DNA isolated from the mouse or hamster donor cell lines. In the somatic cell hybrids, this 249 bp fragment was amplified as expected from DNA samples containing human chromosomes 3, 7 and 19. It was also amplified from DNA samples containing human chromosomes 4, 9, 12, 14, 20, 22 and Y, indicating that at least a small percentage of KRAB-ZFP genes closely related to ZNF91 may be located on these chromosomes.

Detection of sequences homologous to the human ZNF91 KRAB- and spacer-encoding regions in the mouse and rat genomes

To determine whether the mouse and rat genomes contain gene(s) coding for KRAB-ZFPs homologous to the multiple human ZNF91-related genes, we analyzed mouse and rat DNA by Southern blotting with a spacer probe derived from ZNF91 cDNA. This probe is strictly specific to this subset of KRAB-ZFP genes. As a control, the same blot was rehybridized with a KRAB probe. The results shown in Figure 10 indicate that, as previously reported (Bellefroid *et al.*, 1991), the ZNF91 KRAB probe detects multiple fragments in both the rat and mouse genomes. In contrast,

no cross-reacting bands appear under the same hybridization and washing conditions with the ZNF91 spacer probe. These results, in agreement with those obtained by Northern analysis (as outlined below), suggest that no gene homologous to ZNF91 is present in the mouse or rat genome. This could also mean that this specific subset of KRAB-ZFP genes appeared late during evolution.

Expression of ZNF91 and related KRAB-ZFP genes

The expression characteristics of the ZNF91 gene were analyzed by Northern blotting with the ZNF91 spacer probe. This probe was first tested on a Southern blot of human genomic DNA digested with various restriction enzymes. Under high stringency conditions, it was found to detect only one major signal (data not shown). Under identical high stringency conditions, a 4.4 kb transcript was found to hybridize in a variety of cell lines, though at substantially different levels (Figure 11A). The highest relative concentrations were found in T lymphoid cells (Molt4, Molt3, HSB2, CCRF-CEM and CEM-CM3). Moderate amounts of transcript were detected in the THP1 myeloid cell line, the NT2D1 teratocarcinoma cell line, the glioblastoma A172 cell line, the IMR32 neuroblastoma cell line and the JEG3 and BeWo choriocarcinoma cell lines. In overexposed autoradiographs (data not shown), weak signals were detected in all the cell lines. Additional shorter transcripts which may correspond to the ZNF91α and β isoforms were also detected, confirming that these isoforms are expressed at much lower levels. ZNF91 mRNA was also detected. In normal tissues and cell populations (Figure 11B), the ZNF91 gene is abundantly expressed in thymus cells, peripheral blood mononuclear cells and purified T lymphoid cells. Moderate to weak expression was observed in all the other RNA preparations analyzed. In contrast, no signal was detected in the rat and mouse cells and tissues tested (i.e. neurons and astrocytes).

Previous hybridizations with distinct ZNF91-related cDNA probes have shown that the corresponding genes show highly similar expression profiles. Sequence analysis of randomly picked ZNF91 related cDNAs have also shown that the signal observed for each probe was not due to cross-hybridization with one or very few genes expressed at a much higher level than the others (Bellefroid *et al.*, 1991). Therefore, to study expression of the numerous ZNF91-related genes in the same type of experiment, filters were rehybridized with a 249 bp fragment mixture obtained by PCR on human genomic DNA with the degenerate primers that selectively amplify the spacer region of all ZNF91-related genes (see Materials and methods). When this probe was used, the size of the RNA transcripts observed and the relative intensities of the signals generated in the different RNA preparations were nearly identical to those obtained with the ZNF91-specific probe (Figure 11). This suggests that, as a family, all ZNF91-related genes produce RNA transcripts of similar size, again reflecting their highly conserved structural organization. Although widely expressed in most tissues, these ZNF91-related genes appear to be expressed at higher levels in T-lymphoid cells.

Discussion

Our evidence, based on techniques such as *in situ* hybridization to human metaphase chromosomes and the isolation and characterization of YAC clones carrying multiple KRAB-ZFP transcription units, suggests a clustered organization of a subset of human KRAB-ZFP genes. Furthermore, comparative sequence analysis of cDNA and genomic clones demonstrates that these KRAB-ZFP genes have a highly conserved exon/intron organization. Finally, expression studies reveal that these KRAB-ZFP genes, although widely transcribed in all tissues, are most strongly expressed in T-lymphoid cells.

Indications of a clustered organization of ZFP-encoding genes on human and mouse chromosomes have been reported previously. These were based on chromosomal mapping data (Crossley and Little, 1991; Huebner *et al.*, 1991; Hoovers *et al.*, 1992; Lichter *et al.*, 1992; Rousseau-Merck *et al.*, 1992). In none of the above cases, however, was the structure of the putative clusters characterized, nor were correlations between their genomic organization and expression characteristics established.

Here, we have analyzed a set of 26 YACs containing multiple ZNF91-related genes. Twenty-one of them were mapped to chromosome region 19p12-p13.1, indicating that most of these genes are clustered in this region. In the YACs, we have identified 23 distinct ZNF91-related genes. These genes represent only part of the cluster since (i) we have not identified by sequencing all the YAC-borne ZNF91-related genes; (ii) the set of YACs reported here does not encompass the entire segment containing the cluster of ZNF91-related genes (YACs 332C9, 158C9 and 411H1 appear to fall outside any contig); (iii) the ZNF92 gene assigned to 19p12-p13.1 and the ZNF96, 97, 98 and 99 genes derived from the chromosome-19-specific library are very likely to be additional members of this cluster. To construct a contiguous map of YACs spanning the entire cluster, additional YAC clones need to be isolated and characterized by PFGE. The relative order of the different ZNF91-related genes will be determined by isolating and arranging into contigs the corresponding cosmids from the chromosome-19-specific library, using the labeled YACs as probes.

We do not know whether all these genes are transcribed. Some may be cDNA-like pseudogenes as described for the HF.10 gene (Lanfrancone *et al.*, 1992). Other previously characterized KRAB-ZFP genes have been located on other chromosomes: Kox1(ZNF10), 12q12-qter; Kox8(ZNF15), 7q (Huebner *et al.*, 1991); ZNF2, 2; ZNF41, X (Rosati *et al.*, 1991); ZNF7, 8q24; ZNF8, 20q13 (Lania *et al.*, 1990); and ZNF45, 19q (Constantinou-Deltas *et al.*, 1992). Whether these genes unrelated to ZNF91 are also parts of gene clusters and whether these subgroups of structurally related KRAB-ZFP genes are involved in different functional processes is also unknown.

All the KRAB-ZFP genes characterized in this study have exactly the same pattern of exon/intron organization. The conserved KRAB domain is encoded by two separate exons and the variant zinc finger domain is encoded by a single exon, as in most ZFP genes characterized to date except, for example, the TFIIIA (Tso *et al.*, 1986) and WT1 (Gessler *et al.*, 1992) genes. This organization of the ZNF91 gene is consistent with the hypothesis that the distinct mRNA isoforms identified by PCR and most probably representing alternative ZNF91 transcripts are generated by differential splicing. However, in the case of a large family of closely related genes such as the KRAB family, final proof of

differential splicing of the primary transcript of a given gene member will require isolating the corresponding cDNA clones extending into the variable zinc finger domain. Such alternative cDNA forms with variability within the KRAB domain have already been described for two KRAB-ZFP genes: ZNF43, a member of the ZNF91 subgroup (Bellefroid *et al.*, 1991; Lovering and Trowsdale, 1991), and ZNF2 (Rosati *et al.*, 1991). Alternative splicing has also been documented for the *Drosophila* broad complex (Di Bello *et al.*, 1991), *tramtrack* (Read and Manley, 1992) and *WT1* genes (Gessler *et al.*, 1992). In these cases, the genes encode ZFPs presenting variability in the finger domain and thus distinct DNA-binding specificities. In contrast, the variable modular composition of the *Xenopus* ZFPs of the FAX family has been attributed to other molecular mechanisms: the absence of particular modules of the FAX domain is here the result of internal deletions/insertions in exons and mutational activation/inactivation of cryptic splice sites (Nietfeld *et al.*, 1993).

In agreement with previous studies (Bellefroid *et al.*, 1991; Lovering and Trowsdale, 1991), we have shown that the ZNF91-related genes, although transcribed in a variety of tissues, have a common expression profile with increased expression in T lymphoid cells. We do not know whether this enhanced expression of ZNF91-related genes is of any particular importance in this cell type or whether, like members of other multigene families (e.g. HOX class homeobox genes), members of this gene cluster participate in the regulation of a common developmental program. The huge number of ZNF91-related genes also raises the question of the origin and biological significance of this gene redundancy. Expression studies with several probes for individual members of the gene cluster will be necessary to characterize this gene family further. It will be also of interest to determine the location of the ZNF91-related genes with respect to uncharacterized chromosomal translocations such as the t(4;19) (q21;p13) found in acute lymphoblastic leukemias (Saltman *et al.*, 1990). Comparative mapping studies in other species with the YACs as probe could also give information concerning the conservation of this gene cluster that, as suggested by our hybridization data, probably appears late during evolution.

At this stage of the work, it is tempting to speculate that the ZNF91-related ZFPs, by analogy to *Krüppel* and other ZFPs, are developmental regulators. Indeed, some members of the KRAB-ZFP family display expression profiles suggestive of a potential role in the control of development. The murine Zfp2 gene, where the KRAB-homologous domain is surprisingly located in the 5′ untranslated part of the gene, is specifically expressed in the developing nervous system (Chowdhury *et al.*, 1988). The presence of the *Xenopus* Xfin ZFP is apparently associated with the terminal differentiation of cells (De Lucchini *et al.*, 1991).

However, numerous members of the KRAB family (Chowdhury *et al.*, 1989; Lania *et al.*, 1990; Rosati *et al.*, 1991; Thiesen *et al.*, 1991), like a large number of other vertebrate ZFP genes (Köster *et al.*, 1991), have a broad expression profile. Moreover, the *Xenopus* Xfin KRAB-ZFP has been located in the cytoplasm (De Lucchini *et al.*, 1991). This suggests that, as exemplified by TFIIIA, p43 and XFG5-1 ZFPs (Joho *et al.*, 1990; Theunissen *et al.*, 1992; Van Wijck *et al.*, 1992), Xfin may play a regulatory role as an RNA-binding protein. ZNF85 KRAB-ZFP expressed in bacteria has also been shown to bind, at least non-specifically, both DNA and RNA (Marine,J.-C., Bellefroid,E.J. and Martial,J.A., unpublished data). These observations support the hypothesis that at least some KRAB-ZFPs may function through mechanisms other than transcriptional regulation, as exemplified by the *Drosophila Krüppel* ZFP during embryonic development.

Materials and methods

Isolation and characterization of human KRAB-ZFP cDNAs and genomic clones

ZNF91 cDNA was isolated from a human λgt10 cDNA library constructed from RNA isolated from undifferentiated NT2D1 cells (Skowronski *et al.*, 1985) by screening as described in Bellefroid *et al.* (1991). Genomic clones were isolated from a chromosome-19-specific cosmid library in Lawrist 5 (de Jong *et al.*, 1989) by screening with probes derived from ZNF91 cDNA. The fragments used were: N1−160 (5′ region), N31−N335 (KRAB domain), and N372−N693 (spacer region). Probes were ^{32}P-labeled using Amersham's Multiprime DNA labeling system. Hybridizations were performed in 5×SSPE, 5×Denhardt's, 0.1% SDS and denatured salmon sperm DNA (100 μg/ml) at 65°C for 16 h. Filters were washed in either 2×SSC, 0.1% SDS at 50°C (low stringency conditions) or 0.1×SSC, 0.1% SDS at 65°C (high stringency conditions) and subjected to autoradiography at −70°C with Fuji X films. The ZNF91 cDNA was completely sequenced. Cosmids F11465, F21946, F28423, F31854 and F33152 were analyzed by restriction mapping and hybridization with ZNF91 probes representative of the entire mRNA. *Eco*RI and *Hin*dIII DNA fragments were cloned in the pGEM3Zf(−) vector (Promega) and used for subsequent subcloning and DNA sequencing. DNA sequence analysis was performed along the entire exons for F21946 or only at the exon/intron boundaries for the other genomic fragments with the Sequenase version 2 DNA sequencing kit (US Biochemical) and universal M13 or synthetic oligonucleotide primers (Eurogentec). Sequences were compiled with the Sequence Analysis Software Package (version 5) of the University of Wisconsin Genetics Computer Group.

Isolation and characterization of YAC clones.

Clones from the YAC library of the Centre d'Etude du Polymorphisme Humain (CEPH) (Albertsen *et al.*, 1990) were isolated by conventional PCR-based screening (Green and Olson, 1990) with degenerate primers based on a consensus derived from a compilation of nucleotide sequences of the spacer region of ZNF91-related KRAB-ZFPs (Bellefroid *et al.*, 1991). These primers [5′-ATTTT(A/G/C)C(C/T)C(A/G)AGA(T/C)(C/G)TTTG(G/T)CC-(A/G)GAGC and 5′-TGTCT(G/A)TTTGAATTT(G/A)AAA(T/A)T(T/A)TAT (Eurogentec), spanning nucleotides 359−386 and 585−609 (Figure 1)] were used at an annealing temperature of 52°C exactly as described below. PCR products were subcloned into pUC13 plasmids (Pharmacia) and sequenced.

Yeast DNA was dissolved, YAC agarose blocks prepared and restriction enzyme digestions performed as described previously (Albertsen *et al.*, 1990). Pulsed field gel electrophoresis (PFGE) was performed according to the manufacturer's instructions (Biometra). Fingerprinting of the YAC clones by the Alu-PCR method was performed on DNA plugs melted and diluted 1:40 in the reaction mixture (see below). The Alu primer TC-65 (Nelson *et al.*, 1989) was used and incubation conditions were as follows: 1 min at 94°C, 1 min at 46°C and 4 min at 72°C for 35 cycles.

PCR amplification and analysis of PCR products.

For the identification of ZNF91 mRNA isoforms, 20 μg of total RNA was reverse-transcribed into single-stranded DNA at 37°C for 60 min in a 100 μl reaction containing 40 mM KCl, 50 mM Tris−HCl (pH 8.3), 8 mM $MgCl_2$, 0.5 mM dNTPs, 100 ng of random hexanucleotide primers (Eurogentec), 100 units of murine MLV reverse transcriptase and 100 units of RNAsin (BRL). First strand cDNAs were then extracted with phenol−chloroform and precipitated. One-tenth of the cDNA was amplified in a 100 μl PCR reaction mixture containing *Taq* polymerase buffer [50 mM KCl, 10 mM Tris−HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.1% gelatin] supplemented with 100 pmol each of sense (5′-TGCTCCAGGAGGCCC-AGCCTG) and antisense primers (5′-GTTGGGTTTTGTGTAAACGGA-TGCAA) of ZNF91 (N31−N51 and N668−N693), 200 μM each of the four dNTPs, and 2.5 U *Taq* polymerase (Promega). Thirty cycles of amplification were performed with an automated thermal cycler (EquiBio) under the following conditions: 1 min at 94°C; 1.5 min at 65°C; 1 min at 72°C. PCR products were analyzed by 1% agarose gel electrophoresis,

purified on Sephacryl columns (Pharmacia), and directly sequenced, the amplimers being used as primers (Bachmann *et al.*, 1990).

For chromosomal mapping of the ZNF91-related KRAB-ZFP genes, the degenerate oligonucleotides used were the same as for YAC library screening. PCR was carried out, as described above, with 1 μg of human, mouse and rat DNA, and DNA samples from the various monochromosomal somatic cell hybrids of panel no. 2 (Corriel Cell Repositories). An initial 5′ denaturation step at 95°C was performed, followed by 30 cycles of 94°C (1 min), 52°C (1 min) and 72°C (4 min). PCR products were analyzed by Southern analysis as described below with a ^{32}P-labeled ZNF91 cDNA fragment (N372–N693). Hybridization and washing were as described for library screening.

Southern blot analysis, probes and nomenclature

Isolation of total genomic DNA from human, mouse and rat cells and from each of the human–rodent hybrid cell lines was performed as previously described (Wathelet *et al.*, 1988). After electrophoretic separation of restriction fragments on 1% agarose gels and transfer to Hybond N+ membranes (Amersham), hybridization and washing were performed as for library screening (see above).

The KRAB-ZFP-encoding cDNAs described by Bellefroid *et al.* (1991) were subcloned in pGEM3Zf(−) vectors. Complete or partial cDNA inserts were used as ^{32}P-labeled probes. Nomenclature was established and gene symbols assigned by Phyllis J.McAlpine (Chair, HGM nomenclature, University of Manitoba; Winnipeg, Manitoba, Canada). The correspondence between the nomenclature used in Bellefroid *et al.* (1991) and the HGM nomenclature is as follows: ZNF85/HTF1, ZNF43/HTF6, ZNF90/HTF9, ZNF91/HTF10, ZNF92/HTF12 and ZNF93/HTF34. The KRAB and spacer probes used to detect homologous fragments in mouse and rat DNAs were from ZNF91 cDNA (N31–N335 and N372–N693).

Cell lines and preparation of normal cell populations

Cell lines were purchased from the American Type Culture Collection except for NT2D1 (Andrews *et al.*, 1984), LB-33-MEL, LB-37-NSCLC and LB-23-SAR (obtained from the Ludwig Institute, Brussels), and LHN13 (obtained from the Necker Hospital, Paris). The lines were maintained in the recommended media supplemented with 10% fetal calf serum at 37°C and 5% CO_2.

The following human cell lines were used: NT2D1 (testis teratocarcinoma); PA1 (ovarian teratocarcinoma); U87MG and A172 (glioblastomas); IMR32 (neuroblastoma); JEG3 and BeWo (choriocarcinomas); HepG2 (hepatocellular carcinoma); HeLa (cervical carcinoma); OVCAR3 (ovarian carcinoma); MRC5 (fetal lung fibroblast); MeWo and LB-33-MEL (melanomas); LB-37-NSCLC (non small cell lung carcinoma); LB-23-SAR (sarcoma). T-lymphoid lines included: HSB2 (peripheral blood acute lymphoblastic); CCRF-CEM, MOLT3 and MOLT4 (acute lymphoblastic leukemias). B lymphoid lines included: LHN13 (EBV-transformed); IM9 (multiple myeloma) and Daudi (Burkitt lymphoma). Myeloid lines included: U937 (hystiocytic leukemia); KG1 (acute myelogenous leukemia); THP1 (acute monocytic leukemia), and K562 (chronic myelogenous leukemia).

Primary cultures of human smooth muscle cells, skin fibroblasts, umbilical cord endothelial cells and foreskin keratinocytes were kindly provided by Dr C.Lambert (Université de Liège). Primary cultures of rat cerebellar neurons and cerebral cortex astrocytes were obtained from Dr P.Delree (Université de Liège). Cultured cytothrophoblast cells from first trimester or term placenta were provided by Dr E.Alsart (Ecole Normale Supérieure, Paris). Peripheral blood mononuclear cells (PBMC) were isolated from healthy adult donors by Ficoll density gradient centrifugation. T cells were separated from PBMC by rosetting with aminoethylisothiouranium bromide-treated sheep red blood cells (Sigma). This yielded populations that were >85% $CD3^+$, as determined by flow cytometry. B cells were obtained from tonsillar lymphoid cells by panning with anti-IgM (Wisocki and Sato, 1978); this population contains >80% B cells as determined by flow cytometry. Adherent PBMC were purified as described elsewhere (Bergholtz and Thorsby, 1978).

RNA isolation and Northern blot analysis

Total RNAs were isolated by the guanidinium isothiocyanate method (Chirgwin *et al.*, 1979) and equal amounts (10 μg/lane) were size-fractionated by electrophoresis through 1% agarose–formaldehyde gels, transferred to Hybond N membranes (Amersham) and hybridized with randomly primed ^{32}P-labeled probes, exactly as described in Bellefroid *et al.* (1991). The spacer probe was derived from ZNF91 cDNA (N372–N693), the PCR mix probe was a 249 bp fragment obtained by PCR on human total genomic DNA with the degenerate primers used for YAC library screening (see above). Final washing was at 65°C in 0.3×SSC, 0.1% SDS (ZNF91 probe) or at 50°C in 2×SSC, 0.1% SDS (PCR mix probe). Filters were exposed to autoradiographic films at −70°C with intensifying screens for 2–3 days. Gels were stained with ethidium bromide and blots were rehybridized with a 28S rRNA oligonucleotide probe to monitor the amount of RNA loaded (Barbu and Dautry, 1989).

In situ hybridization

Metaphase chromosomes were prepared as described by Ried *et al.* (1992). Briefly, the chromosomes were denatured for 2 min at 80°C in 70% formamide, 2×SSC and dehydrated through an ethanol series (70/90/100%). Probes, either total cDNA fragments of previously described KRAB-ZFPs (Bellefroid *et al.*, 1991) or Alu-PCR products derived from the YACs generated as described by Lengauer *et al.* (1992), were labeled with biotin-11-dUTP (Sigma) in a standard nick translation reaction. The DNase I concentration was adjusted to produce probe fragments of between 150 and 500 bp in size. 50–150 ng of DNA probe was coprecipitated with 5 μg yeast RNA and 5 μg salmon sperm DNA (cDNA probes) or with 5 μg salmon sperm DNA and 50 μg of Cot-1 DNA (YAC Alu-PCR products). The DNA probe was resuspended in 10 μl of 50% formamide, 2×SSC, 10% dextran sulfate. After denaturation (5 min, 76°C), the hybridization solution was applied to the prewarmed slide and overlaid with an 18×18 mm coverslip. For the labeled PCR products, a 30 min preannealing step was performed at 37°C before hybridization. After overnight incubation at 37°C, slides were washed in 50% formamide, 2×SSC (three times 5 min, 42°C) followed by three washes in 0.1×SSC (5 min each, 60°C). After a blocking step in 4×SSC, 3% BSA (37°C, 30 min), the biotinylated probe was detected with avidin-DCS- fluorescein (Vector Laboratories, 5 μg/ml final) 30 min, 37°C, in 4×SSC, 1% BSA. The slides were then washed three times for 5 min at 4°C in 4×SSC, 1% Tween. Chromosomes were counterstained with DAPI and propidium iodide. For experimental details, see Ried *et al.* (1992). Images were produced with a Zeiss Axioskop epifluorescence microscope equipped with a cooled CCD camera (Photometrics PM152, Tucson, AZ). The camera was controlled by and images merged with an Apple Macintosh IIX computer running a software program developed in D.C.Ward's laboratory by Timothy Rand. Images were photographed from the screen on Kodak 100HC color slide films.

Acknowledgements

We thank M.F.Singer and R.Thayer for the gift of the NT2D1 cell cDNA library, D.Cohen and D.Lepaslier for providing the YAC library and for their helpful advice. We are also grateful to N.Antoine, E.Alsart, P.Delree, C.Lambert, P.Lefevre and E.Franzen, for the gift of primary cell cultures, to C.Lacombe and S.Gisslbrecht for providing us with fetal liver RNA, to C.Lengauer and T.Cremer for their help with the Alu-PCR labeling of YAC clones and to J.Lejeune for secretarial assistance. This research was supported by the Belgian National Fund for Scientific Research (FSRM 3.4540.91 and Televie 7.4574.90) and by the Service de la Programmation de la Politique Scientifique (Science de la Vie Bio/15 and PAI.5bis) (grants to J.A.M.); by the Fund for Medical Scientific Research (FRSM, Belgium) and the CGER-ASLK (grants to C.S. and J.S.); and by NIH grants HG00272 to D.C.W. Work performed at Lawrence Livermore National Laboratory was under the auspices of the U.S. Department of Energy, Contract No. W-7405-ENG-48. D.A.P., P.J.L and J.C.M. are fellows of the 'Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture' (IRSIA, Belgium). T.R. received a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. C.S. is Research Director of the National Fund for Scientific Research (FNRS, Belgium).

References

Albertsen,H.M., Abderramin,H., Cann,H.M., Dausset,J., Le Paslier,D. and Cohen,D. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4256–4260.

Andrews,P.W., Damnanov,I., Simon,D., Banting,G., Dracopoli,N.C. and Fogh,J. (1984) *Lab. Invest.*, **50**, 147–162.

Bachmann,B, Lüke,W. and Hunsmann,G. (1990) *Nucleic Acids Res.*, **18**, 1309.

Barbu,V. and Dautry,F. (1989) *Nucleic Acids Res.*, **17**, 7115.

Bellefroid,E.J., Lecocq,P., Benhida,A., Poncelet,D.A., Belayew,A. and Martial,J.A. (1989) *DNA*, **8**, 377–387.

Bellefroid,E.J., Poncelet,D.A., Lecocq,P.J., Revelant,O. and Martial,J.A. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 3608–3612.

Bellefroid,E.J., Ried,T., Riviere,M., Marine,J-C., Levan,G., Szpirer,J., Szpirer,C., Ward,D.C. and Martial,J.A. (1992) *Genomics*, **14**, 780–781.

Bergoltz,B.O. and Thorsby,E. (1978) *J. Immunol. Methods*, **8**, 63.

Bordereaux,D.,Fichelson,S., Tambourin,P. and Gisselbrecht,S. (1990) *Oncogene*, **5**, 925–927.

Chavrier,P., Lemaire,P., Revelant,O., Bravo,R. and Charnay,P. (1988) *Mol. Cell. Biol.*, **8**, 1319–1326.
Chirgwin,J.M., Przybila,A.E., McDonald,R.J. and Rutter, W.J. (1979) *Biochemistry*, **18**, 5294–5299.
Chowdhury,K., Deutsch,V. and Gruss,P. (1987) *Cell*, **48**, 771–778.
Chowdhury,K., Dressler,D., Breier,G., Deutch,U. and Gruss,P. (1988) *EMBO J.*, **7**, 1345–1353.
Chowdhury,K., Dietrich,S., Balling,R., Guenet,J.L. and Gruss,P. (1989) *Nucleic Acids Res.*, **17**, 10427–10438.
Constantinou-Deltas,C.D., Gilbert,J., Bartlett,R.J., Herbstreith,M., Roses,A.D. and Lee,J.E. (1992) *Genomics*, **12**, 581–589.
Crossley,P.H. and Little,P.F.R. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 7923–7927.
Cunliffe,V., Koopman,P., McLaren,A. and Trowsdale,A. (1990) *EMBO J.*, **9**, 197–205.
de Jong,P.J., Yokobata,K., Chen,C., Lohman,F., Pederson,L., McNinch,J. and Van Dilla,M. (1985) *Cyt. Cell Genet.*, **51**, 985.
De Lucchini,S., Rjili,F., Ciliberto,G. and Barsacchi,G. (1991) *Mech. Dev.*, **36**, 31–40.
Di Bello,P.R., Witters,D.A., Bayer,C.A., Fristrom,J.W. and Guild,G.M. (1991) *Genetics*, **129**, 385–397.
El-Baradi,T. and Pieler,T. (1991) *Mech. Dev.*, **35**, 155–169.
Gessler,M., Poustka,A., Cavennee,W., Neve,R.L., Orkin,S.H. and Bruns,G.A.P. (1990) *Nature*, **343**, 774–778.
Green,E.D. and Olson,M.V. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1213–1227.
Haber,D.A., Buckler,A.J., Glaser,T., Call,K.M., Pelletier,J., Sohn,R.L., Dougflass,E.C. and Housman,D.E. (1990) *Cell*, **61**, 1257–1269.
Hoovers,J.M.N., Mannens,M., John,R., Bliek, J., Van Heyningen,V., Porteous,D.J., Leschot,N.J., Westerveld,A. and Little,P.F.R. (1992) *Genomics*, **12**, 254–263.
Huebner,K., Druck,T., Croce,C.M. and Thiesen,H.J. (1991) *Am. J. Hum. Genet.*, **48**, 726–740.
Joho,K.E., Daeby,M.K., Crawford,E.T. and Brown,D.D. (1990) *Cell*, **61**, 293–300.
Knöchel,W., Poting,A., Köster,M., El-Baradi,T., Nietfeld,W., Bouwmeester,T. and Pieler,T. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6097–6100.
Köster,M., Pieler,T., Pöting,A. and Knöchel,W. (1988) *EMBO J.*, **7**, 1735–1741.
Köster,M., Kühn,U., Bouwmeester,T., Nietfeld,W., El-Baradi,T., Knöchel,W. and Pieler,T. (1991) *EMBO J.*, **10**, 3087–3096.
Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8148.
Kroczek,R.A. (1989) *Nucleic Acids Res.*, **17**, 9497.
Lanfrancone,L., Pengue,G., Pandoli,P.P., Salcini,A.E., Giacomucci,A., Longo,L., Donti,E., De Luca,P., La Mantia,G., Pelicci,P.G. and Lania,L. (1992) *Genomics*, **12**, 720–728.
Lania,L., Donti,E., Pannuti,A., Paccucci,A., Pengue,G., Feliciello,I., La Mantia,G., Lanfrancome,L. and Pelicci,P.G. (1990) *Genomics*, **6**, 333–340.
Lengauer,C., Green,E.D. and Cremer,T. (1992) *Genomics*, **13**, 826–828.
Lichter,P., Tang,C.C., Call,K., Hermanson,G., Evans,G.A., Housman,D. and Ward,D.C. (1990) *Science*, **247**, 64–69.
Lichter,P., Bray,P., Ried,T., Dawid,I.B. and Ward,D.C. (1992) *Genomics*, **13**, 999–1007.
Lovering,R. and Trowsdale,J. (1991) *Nucleic Acids Res.*, **19**, 2921–29xx.
Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609–1614.
Nelson,D.L., Ledbetter,S. A., Corbo,L., Victoria,M.F., Amires-Solis,R., Webster,T.D., Ledbetter,D.H., and Caskey,C.T. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6686–6690.
Nietfeld,N., Conrad,S., Van Wijck,I., Giltay,R., Bouwmeester,T., Knöchel,W.,and Pieler,T. (1993) *J. Mol. Biol.*, in press.
Pavletich,N.P. and Pabo,C.0. (1991) *Science*, **252**, 809–816.
Read,D. and Manley,J.L. (1992) *EMBO J.*, **11**, 1035–1044.
Ried,T., Baldini,A., Rand,T.C. and Ward,D.C. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 1388–1392.
Rosati,M., Marino,M., Franze,A., Tramontano,A. and Grimaldi,G. (1991) *Nucleic Acid Res.*, **19**, 5661–5667
Rosenberg,U.B., Schröder,C., Preiss, Kienlin,A., Cote,S., Riede,I. and Jäckle,H. (1986) *Nature*, **319**, 336–339.
Rousseau-Merck,M.F., Tunnaclife,A., Berger,R., Ponder,B.A.J. and Thiesen,H.J. (1992) *Genomics*, **13**, 845–848.
Ruiz i Altaba,A., Perry-O'Keefe,H. and Melton,D.A. (1987) *EMBO J.*, **6**, 3065–3070.
Saltman,D.L., Mellentin,J.D., Smith,S.D. and Cleary,M.L. (1990) *Genes Chromosomes and Cancer*, **2**, 259–265.
Schuh,R., Aicher,W., Gaul,U., Cote,S., Preiss,A., Maier,D., Seifert,E., Nauber,U., Schröder,C., Kemler,R. and Jäckle,H. (1986) *Cell*, **47**, 1025–1032.
Shapiro, M.B. and Senapathy,P. (1987) *Nucleic Acids Res.*, **15**, 7155–7174.
Skowronski,J., Fanning,T.G. and Singer,M.F. (1985) *Mol. Cell. Biol.*, **8**, 1385–1397.
Tautz,D., Lehmann,R.N., Schnürch,H., Schih,R., Siefert,E., Kienlin,A., Jones,K. and Jäckle,H. (1987) *Nature*, **327**, 383–389.
Theunissen,O., Rudt F., Guddat,U., Mentzel,H. and Pieler,T. (1992) *Cell*, **71**, 679–690.
Thiesen,H.J., Bellefroid,E.J., Revelant,O. and Martial,J.A. (1991) *Nucleic Acids Res.*, **19**, 104.
Tso,J.Y., Van den Berg,D. and Korn,L.D. (1986) *Nucleic Acids Res.*, **14**, 2187–2200.
Van Wijk,I., Burfeind J. and Pieler,T. (1992) *Mech. Dev.*, **39**, 63–73.
Vortkamp,A., Gessler,M. and Grzeschik,K.-H. (1991) *Nature*, **352**, 539–540.
Vincent,A., Colot,H.V. and Rosbach,M.(1985) *J. Mol. Biol.*, **186**, 149–166.
Wathelet,M. *et al.* (1988) *Somat. Cell Mol. Genet.*, **14**, 415–426.
Wilkinson,D.C., Bhatt,S., Chavrier,P., Bravo,R. and Charnay,P. (1989) *Nature*, **337**, 461–464.
Wysocki,L.J. and Sato,V.L. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 2844–2848.

Received on January 24, 1992; revised on December 22, 1992